Comparing the Efficiencies of Stochastic Isothermal Molecular Dynamics Methods

Noname manuscript No.
(will be inserted by the editor)

Ben Leimkuhler · Emad Noorizadeh · Oliver Penrose

Comparing the efficiencies of stochastic isothermal molecular dynamics methods

Received: date / Accepted: date

Abstract Molecular dynamics typically incorporates a stochastic-dynamical device, a "thermostat," in order to drive the system to the Gibbs (canonical) distribution at a prescribed temperature. When molecular dynamics is used to compute time-dependent properties, such as autocorrelation functions or diffusion constants, at given temperature, there is a conflict between the need for the thermostat to perturb the time evolution of the system as little as possible and the need to establish equilibrium rapidly. In this article we define a quantity called the "efficiency" of a thermostat which relates the perturbation introduced by the thermostat to the rate of convergence of average kinetic energy to its equilibrium value. We show how to estimate this quantity analytically, carrying out the analysis for several thermostats, including the Nosé-Hoover-Langevin thermostat due to Samoletov *et al* (J. Stat. Phys. **128**, 1321-1336 , 2007) and a generalization of the "stochastic velocity rescaling" method suggested by Bussi *et al* (J. Chem. Phys **126**, 014101, 2007). We find efficiency improvements (proportional to the number of degrees of freedom) for the new schemes compared to Langevin Dynamics. Numerical experiments are presented which precisely confirm our theoretical estimates.

Ben Leimkuhler · Emad Noorizadeh
The Maxwell Institute and School of Mathematics, University of Edinburgh, EH9 3JZ, UK

Oliver Penrose
Department of Mathematics and the Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, UK

1 Introduction

The question considered in this paper is how best to use molecular dynamics to compute time-dependent properties (such as autocorrelation functions) when the temperature rather than the energy is specified. One approach is to perform simulations using straightforward Hamiltonian dynamics, attempting to choose the energy of the system so that the temperature (as determined from the kinetic energy of the particles) has the desired value; this can be difficult to achieve in practice for any individual trajectory, so one would typically need to employ an ensemble of trajectories. Generating a good representative ensemble is itself a challenging task, and microcanonical simulations typically are subject to energy drifts which may distort the statistics [5,6]. To help address these difficulties, as well as to reach the Gibbs distribution more quickly in models with corrugated energy landscapes, various methods ("thermostats") have been devised in which some kind of perturbation is introduced into the dynamics; this perturbation can be thought of as representing the effect of a heat bath at a prescribed temperature. If the purpose is to overcome an inherent lack of ergodicity in the molecular model in order to promote, say, more rapid sampling of configurational states, then a large thermostatting perturbation may be needed. If, however, the goal is to calculate time-dependent properties such as self-diffusion constants from molecular simulation, the perturbation should be small enough so as not to affect seriously the dynamics of the system on short times, while at the same time being able to rapidly drive the system into equilibrium with the thermostat, so as to give results that are effectively independent of the initial conditions with a minimal investment in computing time. It is not obvious *a priori* that it is possible to achieve both of these aims simultaneously, so that confidence in the validity of temporal correlation functions computed from thermostatted simulations is often low [22,12, 21]; nevertheless, these methods are frequently used in practice despite the lack of a solid theoretical foundation (see e.g. [4] for a very recent example).

The purpose of the present paper is to put forward a quantitative criterion, which we term the "efficiency," for determining how well a given thermostatting method can satisfy the two conflicting requirements. Our use of the term "efficiency" in this context is not related to the concept of *sampling efficiency* [23], nor is it similar to the common usage in numerical analysis which relates to accuracy achieved by a method for given computational work. Instead, we use the term here to refer to the

extent to which perturbations of microcanonical dynamics associated to a thermostatting method are directly applied for the purpose at hand (i.e that of driving a given system into thermal equilibrium); *an efficient method will make relatively small perturbations to the dynamical system in achieving the target temperature*. Our use of the term efficiency is in some way similar to that of [8], where it relates to the ability of a thermostat imposed only on the boundary of a molecular model to generate a prescribed distribution in the interior.

A straightforward way of controlling the temperature in simulations is to model the random interaction with the heat bath by a stochastic perturbing force. The simplest such method is the so-called Langevin dynamics [13, 7], in which the modelled system is effectively immersed in a fluid of much smaller and lighter Brownian particles which perturb its motion. It is possible to give a rigorous analysis of the convergence to thermodynamic equilibrium in Langevin dynamics.

An alternative approach was discussed in Samoletov et al. [20] which combines the kinetic energy control technique devised by S. Nosé and W. Hoover [18, 9] with a single stochastic perturbation. We term this method the 'Nosé-Hoover-Langevin thermostat' abbreviated to NHL. Leimkuhler et al [16] have demonstrated that this scheme is ergodic for a harmonic system under a mild non-resonance assumption.

A different method of improving on the Langevin thermostat has been proposed by Bussi et al[1, 2]. They call their method stochastic velocity rescaling. Instead of using independent random forces acting on all the degrees of freedom of the system, they use a single random force which (like the deterministic force in Nosé-Hoover dynamics) acts so as to multiply all the velocities by the same factor at each time step. They give numerical results [2] indicating that (for a system of 108 particles) their modification of Hamiltonian dynamics disturbs the motion of the system less than the corresponding Langevin dynamics when the parameters are chosen so as to drive the system to equilibrium at the same rate. Like the Nosé-Hoover thermostat, this is a 'gentle' thermostat in the sense that the perturbing force acting on each particle in the system acts along the direction of motion of that particle. Hoover[9] has argued that this is a desirable feature in a thermostat.

With the exception of Langevin dynamics, little is known about the theoretical rates of convergence of these schemes to thermodynamic equilibrium. In this article we review Langevin dynamics, Nosé-

Hoover Langevin, and describe a generalized stochastic velocity rescaling method that includes as a special case the method of Bussi et al[1,2]. For each method, we calculate the rate of convergence to equilibrium using certain approximations analogous to the constitutive relations used in hydrodynamics and continuum mechanics; these give us estimates for the convergence of average kinetic energy to the prescribed target temperature, and, in particular, the asymptotic convergence rates near equilibrium conditions. In the course of this we also give the critical choice of damping coefficient for Nosé-Hoover Langevin dynamics which guarantees an optimal convergence behavior near equilibrium. We next calculate, for each method, the rate of accumulation of perturbation error at small times by an exact method based on a Maclaurin expansion. On the principle that the most efficient thermostat is the one that gives rise to the smallest cumulative deviation from Hamiltonian dynamics during the time necessary to bring the system to equilibrium at the thermostat temperature, we define the efficiency of a thermostat as the reciprocal of that deviation. Using our analytical method, we find that, for systems with a large number (n) of degrees of freedom, the NHL thermostat and the generalized stochastic velocity rescaling thermostat are more efficient than Langevin dynamics, by a factor of order n.

2 Molecular Simulation Methods

2.1 Hamiltonian dynamics and phase-space averages

We consider a system obeying classical mechanics, with a Hamiltonian function of the form

$$H(p_1, \ldots, q_n) := \sum_{i=1}^{n} \frac{p_i^2}{2m_i} + V(q_1, \ldots q_n), \tag{1}$$

where n is the number of degrees of freedom, $p_1, \ldots, p_n$ are the momentum coordinates and $q_1, \ldots, q_n$ are the position coordinates. The function V is the potential energy and m_i is the mass associated with the ith degree of freedom and is assumed to be constant. The Hamiltonian equations of motion are:

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i} \quad (i = 1, \ldots n). \tag{2}$$

For future reference, we define the kinetic energy

$$K := \frac{1}{2}\sum_{i=1}^{n} p_i \frac{\partial H}{\partial p_i} = \sum_{i=1}^{n} \frac{p_i^2}{2m_i}. \tag{3}$$

The phase points $(p_1, \ldots, q_n)$ denoting the possible dynamical states of the system lie in a $2n$-dimensional 'phase space' $\Gamma := \mathbf{R}^n \times \mathbf{R}^n$ or (if the particles are confined in a periodic box of side ℓ) $\mathbf{R}^n \times (\mathbf{R}/\ell\mathbf{Z})^n$. Points in phase space will sometimes be denoted x rather than $(p_1, \ldots, q_n)$. The volume element in phase space will sometimes be denoted $\mathrm{d}^{2n}x$ rather than $\mathrm{d}p_1 \ldots \mathrm{d}q_n$.

To model the randomness in the choice of the initial dynamical state and also (for the stochastic thermostats) in the time evolution, we shall treat the time-dependent dynamical state $x(t)$ as a stochastic process — i.e. a family of random variables, parametrized by the non-negative real variable t. Likewise, we treat the time-dependent positions and momenta $p_1(t), \ldots, q_n(t)$ as stochastic processes. Associated with these stochastic processes is a space Ω comprising all the possible trajectories (realizations) of the stochastic process, and a probability measure on Ω describing the probabilities of the various events that can be defined in terms of these trajectories. Expectations with respect to this probability measure (which depends on the way the system or ensemble is started out) will be denoted by $\mathbf{E}\{\cdot\}$.

An important case arises when the system is started out by choosing the initial dynamical state $(p_1(0), \ldots, q_n(0))$ at random according to the Gibbs canonical distribution over Γ, whose density is

$$\rho_{\mathrm{eq}}(x) := \frac{1}{Z(T)} e^{-H(x)/\mathrm{k}T} \tag{4}$$

where T is the temperature, k is Boltzmann's constant and $Z(T)$ is the phase integral

$$Z(T) := \int_{\Gamma} e^{-H(x)/\mathrm{k}T}\, \mathrm{d}^{2n}x. \tag{5}$$

For Hamiltonian dynamics and two of the three thermostats considered here, it turns out that with this initial condition the probability distribution of the random variable $x(t) \in \Gamma$ is also canonical, for each positive value of t. The resulting stochastic process serves as a model for thermal equilibrium at

temperature T. Expectations with respect to the probability measure of this stochastic process will be denoted by $\mathbf{E}_{\rm eq}\{\cdot\}$.

In the following, we shall usually write the random variables $p_1(t), \ldots, q_n(t)$ more concisely as $p_1, \ldots, q_n$, while at the same time writing $\mathbf{E}$ as $\mathbf{E}^t$. Thus, $\mathbf{E}^t\{p_i q_i\}$ means the same thing as $\mathbf{E}\{p_i(t)q_i(t)\}$, and $\mathbf{E}^t_{eq}\{p_i(0)p_i\}$ means the same as $\mathbf{E}_{eq}\{p_i(0)p_i(t)\}$.

2.2 Langevin dynamics

For the Langevin thermostat, a frictional term and a stochastic term are added to the equations of motion (2), so that they are replaced by the system of (Ito) stochastic differential equations[1]

$$\mathrm{d}q_i = \frac{\partial H}{\partial p_i}\,\mathrm{d}t, \tag{6}$$

$$\mathrm{d}p_i = -\frac{\partial H}{\partial q_i}\,\mathrm{d}t - \gamma p_i\,\mathrm{d}t + \sqrt{2\gamma m_i \mathrm{k}T}\,\mathrm{d}W_i, \tag{7}$$

where γ is a positive parameter measuring the strength of the coupling between the system and the thermostat, and $W_1(t), \ldots, W_n(t)$ are n independent Wiener processes (Brownian motions).

For each t, the joint probability density of the random variables $p_1(t), \ldots q_n(t)$ will be denoted by $\rho_t(p_1, \ldots, q_n)$. It satisfies the following Fokker-Planck equation (also known[19] as Kolmogorov's forward equation)

$$\frac{\partial \rho_t}{\partial t} = -\sum_i \frac{\partial}{\partial q_i}\left(\frac{\partial H}{\partial p_i}\rho_t\right) + \sum_i \frac{\partial}{\partial p_i}\left(\left(\frac{\partial H}{\partial q_i} + \gamma p_i\right)\rho_t + \gamma m_i \mathrm{k}T \frac{\partial \rho_t}{\partial p_i}\right). \tag{8}$$

The Gibbs probability density defined in (4) is a stationary solution of (8).

Equation (8) is referred to as a "degenerate diffusion equation," since the differential operator on the right side is elliptic in the momentum variables but not the position variables. Nonetheless, it is possible to demonstrate [17] that this operator is *hypoelliptic* [10,11], which implies regularity in C^∞, uniqueness of the stationary measure, and therefore ergodicity. Moreover, assuming that $V(q_1, \ldots, q_n)$ is smooth it is possible to demonstrate exponentially rapid convergence to the canonical Gibbs distribution (4) [17].

[1] For general information about stochastic differential equations, see ref. [19].

2.3 Stochastic velocity rescaling thermostats

In this method, the equation for $\mathrm{d}q_i$ is again (6)) but the equation for $\mathrm{d}p_i$ is now (generalizing the original proposal of Bussi et al.[1,2])

$$\mathrm{d}p_i = -\frac{\partial H}{\partial q_i}\,\mathrm{d}t - \Psi(K)p_i\,\mathrm{d}t + \sqrt{2\mathrm{k}T\Phi(K)}p_i\mathrm{d}W, \tag{9}$$

where $W(t)$ is a single Wiener process, Φ is an arbitrary positive-valued function and the function Ψ is defined by

$$\Psi(K) := (2K - (1+n)\mathrm{k}T)\Phi(K) - 2\mathrm{k}TK\frac{\mathrm{d}\Phi}{\mathrm{d}K}. \tag{10}$$

This definition is chosen so as to make the Gibbs probability density ρ_{eq} a stationary solution of the Fokker-Planck equation for the SDE system (6), (9). The original proposal of Bussi et al. corresponds to the choice

$$\Phi(K) = \frac{\gamma''}{2K}, \quad \text{so that} \quad \Psi(K) = (1 - (n-1)\mathrm{k}T/2K)\,\gamma'', \tag{11}$$

in which γ'' is a positive constant (they call $1/2\gamma''$ the 'relaxation time').

In this article we will assume the following regarding the function Φ:

$$K\Phi(K) \text{ is bounded as } K \to 0, \tag{12}$$

$$\Phi(K) \text{ grows at most polynomially as } K \to \infty. \tag{13}$$

2.4 The Nosé-Hoover-Langevin thermostat

The NHL thermostat combines a negative feedback control of the kinetic energy with a stochastic perturbation. The SDE system for this thermostat, which involves an additional stochastic process $\xi(t)$, is

$$\mathrm{d}q_i = \frac{\partial H}{\partial p_i}\mathrm{d}t, \quad \mathrm{d}p_i = \left(-\frac{\partial H}{\partial q_i} - \xi p_i\right)\mathrm{d}t \quad (i = 1, \ldots n), \tag{14}$$

$$\mathrm{d}\xi = \frac{1}{\mu}\,(2K - n\mathrm{k}T)\,\mathrm{d}t - \gamma'\xi\mathrm{d}t + \sqrt{\frac{2\gamma'\mathrm{k}T}{\mu}}\mathrm{d}W, \tag{15}$$

where γ' and μ are positive constants and $W(t)$ is a single Wiener process.

It can be checked that the augmented Gibbs density

$$\rho'_{\rm eq}(x,\xi) := \frac{1}{Z(T)}\sqrt{\frac{\mu}{2\pi {\rm k}T}}e^{-H(x)/{\rm k}T-\mu\,\xi^2/2{\rm k}T} \tag{16}$$

is a stationary solution of the Fokker-Planck equation for the NHL system (14), (15). This distribution therefore takes the place of the Gibbs canonical distribution in the description of equilibrium for the NHL thermostat. It is known [16] that the scheme is ergodic for the case of a harmonic system with a mild assumption on its spectrum.

3 The rate of convergence to the thermostat temperature

The point of using any of these thermostats is to get from an arbitrary initial dynamical state or probability distribution, which may have the wrong energy for the prescribed temperature T, to a state or probability distribution with the right energy. In this section we estimate the length of time this takes, by estimating how long the energy takes to get from an arbitrary initial value to its equilibrium value at the thermostat temperature T. For simplicity the calculations will be done for a system of particles all having the same (constant) mass m.

3.1 The convergence rate for Langevin dynamics

Consider a system which is started in equilibrium at a given temperature T_0 different from the thermostat temperature T and then evolves according to Langevin dynamics; that is to say, the initial phase point is chosen at random from the Gibbs distribution (4) with $T = T_0$, and the subsequent time evolution is given by the equations of the Langevin thermostat, (6) and (7), with a value for T which is different from T_0.

As a simple criterion for convergence to equilibrium with the thermostat we consider the time evolution of the expectation of the energy, $\mathbf{E}^t\{H\}$ where H is the random variable defined in terms of $p_1,\dots q_n$ using the formula (1). By the Ito-Doeblin formula (Theorems 4.1.2 and 4.2.1 in ref. [19]), it

follows from equations (6) and (7) that

$$\mathrm{d}H = \sum_i \frac{\partial H}{\partial p_i}\left(-\gamma p_i\,\mathrm{d}t + \sqrt{2\gamma m_i \mathrm{k}T}\mathrm{d}W_i\right) + n\mathrm{k}T\gamma\mathrm{d}t. \tag{17}$$

Taking the expectation of (17), dividing by $\mathrm{d}t$ and using the formula (3), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}^t\{H\} = \gamma[n\mathrm{k}T - \mathbf{E}^t\{2K\}]. \tag{18}$$

To use this differential equation we need a constitutive relation connecting, at any time t, $\mathbf{E}^t\{H\}$ and $\mathbf{E}^t\{2K\}$. To formulate such a relation, we shall make the following assumption, which has something in common with the constitutive relations used in hydrodynamics and continuum mechanics:

Assumption 1 *There is a time-dependent 'empirical temperature' $\theta(t)$ such that the expectation at time t of any phase space function that is even in the momenta is approximately the same as it would be in a canonical distribution at the temperature $\theta := \theta(t)$.*

As a formula, this assumption asserts that there is a $\theta(t)$ such that, for all phase-space functions f that are even in the momentum variables,

$$\mathbf{E}^t\{f(p_1,\ldots,q_n)\} \approx \frac{1}{Z(\theta)}\int_\Gamma \mathrm{d}^{2n}x\, f(x)\mathrm{e}^{-H(x)/\mathrm{k}\theta}, \tag{19}$$

where

$$Z(\theta) := \int_\Gamma \mathrm{d}^{2n}x\,\mathrm{e}^{-H(x)/\mathrm{k}\theta}. \tag{20}$$

In particular, taking f to be the kinetic energy, (19) gives

$$\mathbf{E}^t\{K\} = \tfrac{1}{2}n\mathrm{k}\theta. \tag{21}$$

Exceptionally, we write this application of (19) as an exact equality, since it will be used as our definition of θ. From this definition it follows that $\theta(0) = \mathbf{E}^{t=0}\{2K\}/n\mathrm{k} = T_0$.

If instead we take f to be the total energy, (19) gives

$$\mathbf{E}^t\{H\} \approx U(\theta) := \frac{1}{Z(\theta)}\int_\Gamma \mathrm{d}^{2n}x\, H(x)\mathrm{e}^{-H(x)/\mathrm{k}\theta}. \tag{22}$$

Using (21) and (22) we can bring (18) to the form

$$C(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} = nk\gamma(T-\theta), \tag{23}$$

where $C(\theta)$ is the heat capacity at temperature θ, defined by

$$C(\theta) := \mathrm{d}U(\theta)/\mathrm{d}\theta. \tag{24}$$

It can be verified, using the definition of U in (22), that $C(\theta)$ is positive.

The differential equation (23) has an equilibrium point, obviously stable, at $\theta = T$. An estimate of the rate of convergence to this equilibrium point can be obtained by considering the linearized version of (23), whose general solution is

$$\theta = T + \mathrm{const} \times \mathrm{e}^{-n\mathrm{k}\gamma t/C(T)}; \tag{25}$$

thus we can estimate the rate of convergence to the thermostat temperature as $n\mathrm{k}\gamma/C(T)$. This is the convergence rate that has been entered into Table I.

3.2 Convergence rate for stochastic velocity rescaling

The equations for this method are (6) and (9), with Ψ defined by (10). The Ito-Doeblin formula now gives

$$\mathrm{d}H = \sum_i \frac{\partial H}{\partial p_i}\left(-\Psi(K)p_i\,\mathrm{d}t + \sqrt{2\mathrm{k}T\Phi(K)}\,p_i\,\mathrm{d}W\right) + 2\mathrm{k}T\Phi(K)K\mathrm{d}t, \tag{26}$$

so that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}^t\{H\} &= \mathbf{E}^t\{2K(\mathrm{k}T\Phi(K) - \Psi(K))\} \\ &= \mathbf{E}^t\left\{2K\left([(2+n)\mathrm{k}T - 2K]\Phi(K) + 2\mathrm{k}TK\Phi'(K)\right\}\right.. \end{aligned} \tag{27}$$

where $\Phi'(K) := \mathrm{d}\Phi/\mathrm{d}K$

According to Assumption 1 (equation (19)), we may approximate the left side of (27) by $C(\theta)\,\mathrm{d}\theta/\mathrm{d}t$ and the right side by its expectation under the Gibbs canonical measure at temperature θ. The probability density for K under this measure is proportional to $\mathrm{e}^{-K/\mathrm{k}\theta}K^{n/2-1}$, so that the expectation of the right-hand side of (27) is

$$\frac{1}{Z_K(\theta)}\int_0^\infty 2K\left([(2+n)\mathrm{k}T-2K]\Phi(K)+2\mathrm{k}TK\Phi'(K)\right)\mathrm{e}^{-K/\mathrm{k}\theta}K^{n/2-1}\mathrm{d}K, \tag{28}$$

where

$$Z_K(\theta) := \int_0^\infty \mathrm{e}^{-K/\mathrm{k}\theta}K^{n/2-1}\mathrm{d}K. \tag{29}$$

Assuming that $\Phi(K)$ satisfies the conditions (12)-(13), partial integration gives

$$\begin{aligned}
&\int_0^\infty K^2\Phi'(K)\mathrm{e}^{-K/\mathrm{k}\theta}K^{n/2-1}\mathrm{d}K\\
&= -\int_0^\infty \Phi(K)\frac{\mathrm{d}}{\mathrm{d}K}\left(\mathrm{e}^{-K/\mathrm{k}\theta}K^2K^{n/2-1}\right)\mathrm{d}K\\
&= -\int_0^\infty \Phi(K)\mathrm{e}^{-K/\mathrm{k}\theta}\left(-\frac{1}{\mathrm{k}\theta}K^{n/2+1}+(\tfrac{1}{2}n+1)K^{n/2}\right)\mathrm{d}K.
\end{aligned} \tag{30}$$

With the help of this result the integral in (28) simplifies to

$$4\left(\frac{T}{\theta}-1\right)\int_0^\infty \Phi(K)\mathrm{e}^{-K/\mathrm{k}\theta}K^{n/2+1}\mathrm{d}K, \tag{31}$$

so that our approximate version of (27) can be written

$$\begin{aligned}
C(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} &\approx 4\left(\frac{T}{\theta}-1\right)\frac{\int_0^\infty \Phi(K)\mathrm{e}^{-K/\mathrm{k}\theta}K^{n/2+1}\mathrm{d}K}{Z_K(\theta)}\\
&= 4\,\frac{\int \mathrm{d}p_1\ldots\mathrm{d}q_n K^2\Phi(K)\mathrm{e}^{-K/\mathrm{k}\theta}}{\int \mathrm{d}p_1\ldots\mathrm{d}q_n\mathrm{e}^{-K/\mathrm{k}\theta}}\left(\frac{T}{\theta}-1\right).
\end{aligned} \tag{32}$$

Equation (32) has a stable equilibrium point at $\theta = T$. The approximating linearized equation is

$$C(T)\frac{\mathrm{d}\theta}{\mathrm{d}t} = 4\,\mathbf{E}_{\mathrm{eq}}\{K^2\Phi(K)\}\frac{T-\theta}{T}, \tag{33}$$

whose solutions converge to equilibrium at the rate $4\mathbf{E}_{\mathrm{eq}}\{K^2\Phi(K)\}/TC(T)$. This is the rate of convergence recorded in Table I. For the choice $\Phi(K) = \gamma''/2K$ used in (11), eqn (32) takes the particularly

simple form

$$C(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} \approx \gamma'' n\mathrm{k}\theta\left(\frac{T}{\theta}-1\right), \tag{34}$$

with convergence rate $\gamma'' n\mathrm{k}/C(T)$

3.3 Convergence to the thermostat temperature for the NHL thermostat

Again we consider a system whose initial dynamical state $x(0)$ is chosen at random from a canonical phase-space probability distribution at temperature T_0, but now, in addition, an arbitrary initial value for $\xi(0)$ is prescribed, and the subsequent evolution is determined by the NHL equations (14), (15). The Ito-Doeblin formula now gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}^t\{H\} &= \sum_i \mathbf{E}^t\left\{\frac{\partial H}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t}+\frac{\partial H}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t}\right\}\\ &= -\sum_i \mathbf{E}^t\left\{\frac{\partial H}{\partial p_i}p_i\xi\right\} = -\mathbf{E}^t\{2K\xi\}.\end{aligned} \tag{35}$$

The left-hand side can be approximated using Assumption 1, just as in (23) and (32), but for the right-hand side we need an additional assumption:

Assumption 2

The random variables $\xi(t)$ and $K(t)$, which are obviously uncorrelated when $t = 0$, remain uncorrelated for all $t > 0$:

$$\mathbf{E}^t\{\xi K\} \approx \mathbf{E}^t\{\xi\}\mathbf{E}^t\{K\}. \tag{36}$$

A test of Assumptions 1 and 2, described later in the paper (see Figures 1 and 2), indicates that both of them are reasonably accurate approximations.

Using (19) to evaluate the left side of (35), and (36) followed by (21) for the right side, we obtain

$$C(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}t} \approx -\mathbf{E}^t\{2K\}\mathbf{E}^t\{\xi\} = -n\mathrm{k}\theta\mathbf{E}^t\{\xi\}, \tag{37}$$

where $C(\theta)$ is the heat capacity, defined in (24). To obtain information about $\mathbf{E}^t\{\xi\}$ we take the expectation of (15) and use the definition (21), obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}^t\{\xi\} = \frac{1}{\mu}\mathbf{E}^t\left\{2K - n\mathrm{k}T\right\} - \gamma'\mathbf{E}^t\{\xi\} = \frac{n\mathrm{k}}{\mu}\left(\theta - T\right) - \gamma'\mathbf{E}^t\{\xi\}. \tag{38}$$

Equations (37) and (38) constitute a dynamical system for the two variables $\mathbf{E}^t\{\xi\}$ and θ, which can be studied using the phase plane method. This system has a unique equilibrium point at $\mathbf{E}^t\{\xi\} = 0, \theta(t) = T$. The equilibrium point is stable; one way to see this is to consider the Lyapunov function

$$\Lambda(\mathbf{E}^t\{\xi\}, \theta) := \tfrac{1}{2}(\mathbf{E}^t\{\xi\})^2 + \int_T^\theta \frac{C(\theta_1)(\theta_1 - T)}{\mu\theta_1}\mathrm{d}\theta_1\,, \tag{39}$$

which is a measure of the distance of the phase point $(\mathbf{E}^t\{\xi\}, \theta)$ from the equilibrium point in the phase plane. The time derivative of Λ equals $-\gamma'(\mathbf{E}^t\{\xi\})^2$, and hence Λ decreases monotonically (strictly so, except at those instants when $\mathbf{E}^t\{\xi\} = 0$) and the phase point moves closer and closer to the equilibrium point. A numerical study illustrating this approach to equilibrium, and also indicating that the system (37), (38) gives a reasonably accurate approximation, is given in section 6.2.2 below.

An analytic estimate of the rate of approach to equilibrium can be obtained by linearizing the system (37), (38). After eliminating $\mathbf{E}^t\{\xi\}$ the linearized equations reduce to

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \gamma'\frac{\mathrm{d}\theta}{\mathrm{d}t} + \frac{n^2\mathrm{k}^2T}{\mu C(T)}(\theta - T) = 0. \tag{40}$$

This is the equation of a damped harmonic oscillator. If γ' is less than the critical damping value

$$\gamma'_{\mathrm{crit}} := 2\sqrt{\frac{n^2\mathrm{k}^2T}{\mu C(T)}}, \tag{41}$$

the amplitude of the oscillations dies out exponentially like $\mathrm{e}^{-0.5\gamma' t}$ and therefore more slowly than $\mathrm{e}^{-0.5\gamma'_{\mathrm{crit}}t}$; on the other hand if $\gamma' > \gamma'_{\mathrm{crit}}$ there are no oscillations but the amplitude again dies out more slowly than than $\mathrm{e}^{-0.5\gamma'_{\mathrm{crit}}t}$. Thus, for a given γ'_{crit}, the most rapid convergence to the thermostat temperature is obtained by choosing $\gamma' = \gamma'_{\mathrm{crit}}$. This convergence rate, $0.5\gamma'_{\mathrm{crit}}$, has been entered in Table I.

If γ' is chosen to be zero, which means using the deterministic Nosé-Hoover thermostat rather than NHL, our method predicts that the oscillations will continue for ever, with whatever amplitude they were given initially. This possibility suggests that the deterministic Nosé-Hoover thermostat may be non-ergodic. Other authors [15, 14, 3] have also drawn attention to this possibility.

4 The rate of accumulation of the thermostat perturbation

In each of the thermostats described above at least one extra term is inserted into the Hamiltonian equation of motion $\mathrm{d}p_i/\mathrm{d}t = \partial H/\partial q_i$. Over a period of time the extra term or terms will take the motion of the perturbed system further and further away from the motion of the Hamiltonian system it is meant to approximate, making the methods potentially unreliable for the calculation of properties of the system, such as time-dependent correlation functions, which are determined by the Hamiltonian trajectories.

In this section we estimate the rate at which the resulting perturbation of the motion builds up over time, with a view to comparing this rate with the rate at which the thermostat brings the system to thermal equilibrium, as calculated in the previous section. As a concrete example, we shall calculate how much this perturbation affects the velocity autocorrelation function (VAF) at small values of t.

Assuming for simplicity that all the particles have the same mass m, the VAF can be defined as

$$F(t) := \frac{1}{nm\mathrm{k}T} \sum_{i=1}^{n} \mathbf{E}_{\mathrm{eq}}\{p_i(0)p_i(t)\} = \frac{1}{nm\mathrm{k}T} \sum \mathbf{E}^t_{\mathrm{eq}}\{p_i(0)p_i\}. \tag{42}$$

The prefactor $1/nm\mathrm{k}T$ has been chosen so as to make

$$F(0) = 1 \tag{43}$$

In this formula, the evolution of p is defined by Hamiltonian dynamics. We shall estimate $F(t)$ for small times using Maclaurin's expansion, under the assumption that one of the thermostatted dynamics has been used instead of Hamiltonian dynamics. For this it will be necessary to calculate the first derivative dF/dt at $t = 0$ and, in the case of the NHL thermostat, the second derivative as well.

4.1 Langevin dynamics

Multiplying eqn (7) by $p_i(0)$, taking the expectation with respect to the equilibrium measure and dividing by dt we obtain, since $p_i(0)$ and $\mathrm{d}W(t)$ are statistically independent if $t > 0$,

$$\mathbf{E}^t_{\mathrm{eq}}\left\{p_i(0)\frac{\mathrm{d}p_i}{\mathrm{d}t}\right\} = \mathbf{E}^t_{\mathrm{eq}}\left\{p_i(0)\left[-\frac{\partial H}{\partial q_i} - \gamma p_i\right]\right\} \quad (t > 0) \tag{44}$$

Substituting into (42) and taking the limit $t \searrow 0$ we obtain

$$\left.\frac{\mathrm{d}F_\gamma(t)}{\mathrm{d}t}\right|_{t=0+} = -\frac{\gamma}{nm\mathrm{k}T}\sum_{i=1}^{n}\mathbf{E}_{\mathrm{eq}}\{p_i(0)^2\} = -\gamma \tag{45}$$

so that (having regard to (43))

$$F_\gamma(t) = 1 - \gamma t + O(t^2) \quad (t > 0) \tag{46}$$

where the dependence of F on the Langevin parameter γ is now shown explicitly. Since Hamiltonian dynamics is the same as Langevin dynamics with $\gamma = 0$, the error in $F(t)$ due to the use of Langevin rather than Hamiltonian dynamics is

$$\Delta_{LD}F(t) := F_\gamma(t) - F_0(t) = -\gamma t + O(t^2) \quad (t > 0) \tag{47}$$

Thus for small t the magnitude of the error is γt. This result has been entered into column 3 of Table I.

4.2 Stochastic velocity rescaling

Multiplying eqn (9) by $p_i(0)$, taking the equilibrium expectation and dividing by dt we obtain

$$\mathbf{E}^t_{\mathrm{eq}}\left\{p_i(0)\frac{\mathrm{d}p_i}{\mathrm{d}t}\right\} = \mathbf{E}^t_{\mathrm{eq}}\left\{p_i(0)\left[-\frac{\partial H}{\partial q_i} - \Psi(K)p_i\right]\right\} \quad (t > 0) \tag{48}$$

Substituting into (42) and taking the limit $t \searrow 0$ we obtain the following equation for $F = F_\Phi$:

$$\left.\frac{\mathrm{d}F_\Phi(t)}{\mathrm{d}t}\right|_{t=0+} = -\frac{1}{nm\mathrm{k}T}\sum_{i=1}^{n}\mathbf{E}_{\mathrm{eq}}^{t=0+}\{\Psi(K)p_i(0)^2\} = -\frac{2}{n\mathrm{k}T}\mathbf{E}_{\mathrm{eq}}\{K\Psi(K)\} \tag{49}$$

Using the definition (10) of Ψ and then simplifying the resulting expression by means of integration by parts, as in (30), it can be shown that $\mathbf{E}_{\mathrm{eq}}\{K\Psi(K)\}) = \mathrm{k}T\mathbf{E}_{\mathrm{eq}}\{K\Phi(K)\}$; therefore, since this thermostat reduces to Hamiltonian dynamics in the case $\Phi = 0$, the error in $F(t)$ for small t is

$$\Delta_{VR}F(t) := F_\Phi - F_0 = -\frac{2}{n}\mathbf{E}_{\mathrm{eq}}\{K\Phi(K)\}t + O(t^2) \quad (t > 0). \tag{50}$$

The magnitude of this expression has been entered into column 3 of Table I.

4.3 The NHL thermostat

The error introduced by this thermostat depends on the initial value of ξ. In the following calculation we assume that both this intial value and the initial dynamical state of the system are chosen at random using the augmented Gibbs distribution defined in Equation (16) .

Multiplying the second equation in (14) by $p_i(0)$, taking the equilibrium expectation and dividing by dt we obtain

$$\mathbf{E}_{\mathrm{eq}}^t\left\{p_i(0)\frac{\mathrm{d}p_i}{\mathrm{d}t}\right\} = \mathbf{E}_{\mathrm{eq}}^t\left\{p_i(0)\left[-\frac{\partial H}{\partial q_i} - \xi p_i\right]\right\} \quad (t > 0) \tag{51}$$

Substituting into (42) and taking the limit $t \searrow 0$ we find that F (for given μ) satisfies

$$\left.\frac{\mathrm{d}F_\mu(t)}{\mathrm{d}t}\right|_{t=0+} = -\frac{1}{nm\mathrm{k}T}\sum_{i=1}^{n}\mathbf{E}_{\mathrm{eq}}\{\xi(0)p_i(0)^2\} = 0 \tag{52}$$

because the augmented Gibbs density (16) is an even function of ξ. Thus, the Maclaurin series for $F(t)$ begins with a quadratic term, for which we need the second derivative of $F(t)$ at $t = 0$.

The equation for $\mathrm{d}p_i$ in (14) can be written $\mathrm{d}p_i = y_i\mathrm{d}t$, where

$$y_i := -\frac{\partial H}{\partial q_i} - \xi p_i \tag{53}$$

Using the Ito-Doeblin formula and then the equations (14), (15) which define this thermostat, we obtain

$$\begin{aligned} \mathrm{d}y_i &= -\sum_j \frac{\partial^2 H}{\partial q_i \partial q_j}\mathrm{d}q_j - \xi \mathrm{d}p_i - p_i \mathrm{d}\xi \\ &= -\sum_j \frac{\partial^2 H}{\partial q_i \partial q_j}\frac{\partial H}{\partial p_j} + \xi\left(\frac{\partial H}{\partial q_i} + \xi p_i\right)\mathrm{d}t - \\ &\quad -p_i\left(\frac{1}{\mu}\left(2K - n\mathrm{k}T\right)\mathrm{d}t - \gamma'\xi \mathrm{d}t + \sqrt{\frac{2\gamma' \mathrm{k}T}{\mu}}\mathrm{d}W\right) \quad (t>0) \end{aligned} \tag{54}$$

so that, omitting terms which turn out to be zero,

$$\begin{aligned} &\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathbf{E}_{\mathrm{eq}}\{p_i(0)p_i(t)\}\Big|_{t=0+} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}_{\mathrm{eq}}\{p_i(0)y_i(t)\}\Big|_{t=0+} \\ &= -\mathbf{E}_{\mathrm{eq}}\left\{p_i(0)\sum_j \frac{\partial^2 H}{\partial q_i \partial q_j}\frac{\partial H}{\partial p_j}\right\}\Big|_{t=0+} + \mathbf{E}_{\mathrm{eq}}\{p_i(0)^2\xi(0)^2\} - \mathbf{E}_{\mathrm{eq}}\{p_i(0)^2\frac{1}{\mu}\left(2K|_{t=0+} - n\mathrm{k}T\right)\} \\ &= -\mathbf{E}_{\mathrm{eq}}\left\{p_i(0)\sum_j \frac{\partial^2 H}{\partial q_i \partial q_j}\frac{\partial H}{\partial p_j}\right\}\Big|_{t=0+} + \frac{m(\mathrm{k}T)^2}{\mu} - \frac{2m(\mathrm{k}T)^2}{\mu} \end{aligned} \tag{55}$$

This thermostat reduces to Hamiltonian dynamics in the limit $\mu \to \infty$; hence, using Maclaurin's expansion in the formula (42), we obtain (with the help of (43) and (52))

$$\Delta_{NHL}F(t) = F_\mu(t) - F_\infty(t) = -\frac{1}{nm\mathrm{k}T}\sum_{i=1}^{n}\frac{m(\mathrm{k}T)^2}{\mu}\frac{t^2}{2} + O(t^3) = -\frac{\mathrm{k}T}{\mu}\frac{t^2}{2} + O(t^3) \tag{56}$$

Thus, the error after a small time t is roughly $-\frac{\mathrm{k}T}{2\mu}t^2$. The magnitude of this expression has been entered into column 3 of Table I. It is very interesting to note that, at least for modest times, the growth of perturbation depends only on the coupling parameter μ and not on γ', unless these are directly coupled by the choice to work with critical damping.

5 The efficiency of a thermostat

Our analytic results are summarized in the first three columns of Table I. To define a numerical measure of the efficiency of a given thermostat, consider the amount of error that accumulates during the time that the the system is brought to equlibrium at the thermostat temperature. The larger this amount of

error, the less efficient the thermostat; so we define the efficiency as the reciprocal of the amount of error that accumulates during that time. For the Langevin and stochastic velocity rescaling thermostats, the error accumulates linearly, and so we can take the efficiency to be the rate of convergence to equilibrium divided by the rate of error accumulation. This ratio is the entry in the last column of Table I. For the NHL thermostat, the error accumulates quadratically. If the parameter γ' is given its optimal value, i.e. the critical damping value shown in (41), the rate of decay to equlibrium is $0.5\gamma'_{crit}$ and so the time to reach equilibrium is of order $2/\gamma'_{crit} = \sqrt{\mu C(T)/n^2\mathrm{k}^2T}$. The amount of relative error in $F(t)$ that builds up in this time is, by the result (56), roughly

$$\frac{\mathrm{k}T}{\mu}\frac{1}{2}\left(\frac{2}{\gamma_{crit}}\right)^2 = \frac{\mathrm{k}T}{2\mu}\frac{\mu C(T)}{n^2\mathrm{k}^2T} = \frac{C(T)}{2n^2\mathrm{k}}. \tag{57}$$

Our estimate of efficiency at critical damping is the reciprocal of this quantity, as shown in Table I.

Table I. Estimated rates of convergence to the thermostat temperature, rates of error accumulation, and efficiencies for various thermostats.

thermostat	convergence rate	error (to lowest order in t)	efficiency
Langevin	$\frac{n\mathrm{k}\gamma}{C(T)}$	γt	$\frac{n\mathrm{k}}{C(T)} \approx 1$
velocity rescaling	$4\,\frac{\mathbf{E}_{\mathrm{eq}}\{K^2\Phi(K)\}}{TC(T)}$	$\frac{2}{n}\mathbf{E}_{\mathrm{eq}}\{K\Phi(K)\}\,t$	$O(n)$
VR with $\Phi(K) = \frac{\gamma''}{2K}$	$\frac{n\mathrm{k}\gamma''}{C(T)}$	$\frac{\gamma''}{n}t$	$\frac{n^2\mathrm{k}}{C(T)} \approx n$
NHL	$\frac{\gamma'_{crit}}{2} = \sqrt{\frac{n^2\mathrm{k}^2T}{\mu C(T)}}$ (at critical damping)	$\frac{\mathrm{k}T}{2\mu}t^2$	$\frac{2n^2\mathrm{k}}{C(T)} \approx 2n$

The entry in the final column is the reciprocal of the number obtained from the entry in the previous column by setting t equal to the reciprocal of the entry in the first column (i.e. to our estimate of the time taken for the system to come to equilibrium with the thermostat.) To estimate $C(T)$ we used Dulong and Petit's experimental law $C(T) \approx n\mathrm{k}$

6 Numerical experiments

In this section we we compare our theoretical predictions about convergence and error build-up with numerical simulation of a two-dimensional system comprising 108 particles in a square periodic box interacting via the Lennard-Jones potential

$$V = \sum \varphi_{\mathrm{LJ}}(r_{ij}) := \sum 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right], \tag{58}$$

where ϵ, σ are positive parameters, r_{ij} denotes the distance between the ith and jth particle and the sum goes over all pairs of particles. In theoretical work with the Lennard-Jones potential it is customary to state the results in terms of the reduced temperature and reduced density, defined (for a two-dimensional system) by

$$T^* := \frac{kT}{\epsilon}, \quad \rho^* := \frac{N}{L^2}\sigma^2.$$

where N is the number of particles and L is the side length of the periodic box. In our simulations we used the values $N = 108$ and $\rho^* = 0.86$, and various values for T*.

For Langevin dynamics we used a weak second-order method. The equations of the velocity rescaling thermostat, (6) and (9), are more difficult to integrate accurately, because of the multiplicative noise (i.e. the fact that the coefficient of $\mathrm{d}W(t)$ is not constant). We first split the equations into the deterministic (Hamiltonian dynamics) and stochastic parts. For the deterministic part, we used a Verlet method and for the stochastic part an Euler-Muruyama method [19]. The choice (11) was used for the functions Ψ and Φ, with various values of γ''.

For the NHL thermostat (Equations (14), (15)) we used the following numerical integrator:

$$\begin{aligned}
\hat{\xi}^{k+1/2} &= e^{-\gamma' \Delta t/2}\xi^k + \sqrt{\frac{kT}{\mu}\,(1 - e^{\gamma' \Delta t})}\,\eta^k \\
p^{k+1/2} &= p^k - \frac{\Delta t}{2}\frac{\partial V}{\partial q}(q^k) - \frac{\Delta t}{2}\hat{\xi}^{k+1/2}p^{k+1/2} \\
\xi^{k+1/2} &= \hat{\xi}^{k+1/2} + \Delta t\frac{2K(p^{k+1/2}) - nkT}{\mu} \\
q^{k+1} &= q^k + \Delta t M^{-1}p^{k+1/2} \\
p^{k+1} &= p^{k+1/2} - \frac{\Delta t}{2}\frac{\partial V}{\partial q}(q^{k+1}) - \frac{\Delta t}{2}\xi^{k+1/2}p^{k+1/2} \\
\xi^{k+1} &= e^{-\gamma' \Delta t/2}\xi^{k+1/2} + \sqrt{\frac{kT}{\mu}\,(1 - e^{\gamma' \Delta t})}\,\zeta^k
\end{aligned}$$

where η^k and ζ^k are standard normal random variables. It can be shown that the above method is second-order (in the weak sense). We used step size $\Delta t = 0.01$ for most simulations, but a much smaller step size ($\Delta t = 0.001$) was needed to examine the error growth in the autocorrelation functions using the two gentle thermostats.

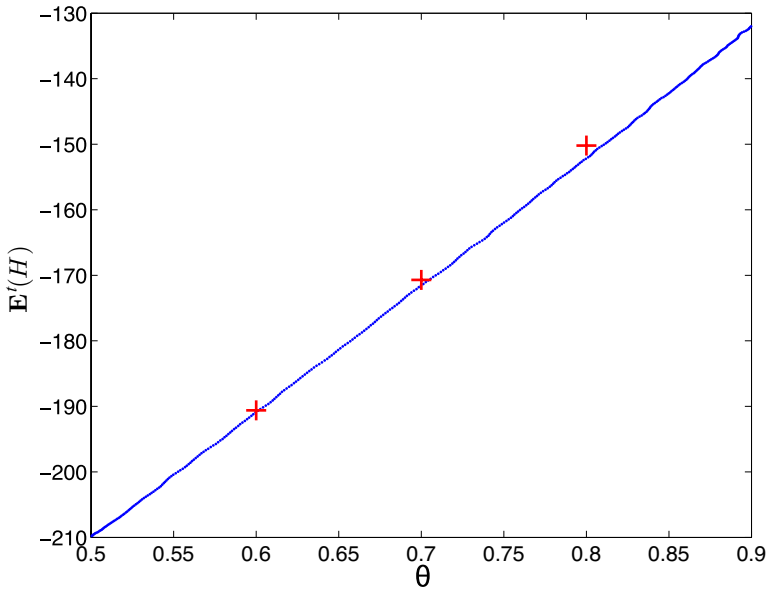

Fig. 1 Computed values of $\mathbf{E}^t\{H\}$ plotted against those of θ, for the simulation of a Lennard-Jones system as described in the text. The axes are labelled with values of the corresponding 'reduced' quantities, *i.e.* $H^* := H/\epsilon$ and $\theta^* := \mathrm{k}\theta/\epsilon$. Ensemble averaged energies are shown by crosses at reduced temperatures $T^* = 0.6, 0.7, 0.8$. (color online)

6.1 Testing the assumptions used in Section 3

The theory of convergence to the thermostat temperature presented in Section 3 depends on Assumptions 1 and 2. We checked the validity of these assumptions using a sample of 10,000 trajectories calculated with the NHL integrator described above. The initial states of the system were chosen at random using the Gibbs probability distribution (4) at (reduced) temperature $T_0^* = 0.5$. The initial value of ξ was 0 for all trajectories. The reduced temperature of the thermostat was $T^* = 1$.

Assumption 1 implies (see Equation (22)) that the graph of $\mathbf{E}^t\{H\}$ against θ should be the same as the graph of the thermodynamic internal energy $U(T)$ against temperature T, whose slope is equal to the heat capacity. To test this, we computed estimates of $\mathbf{E}^t\{H\}$ and θ by averaging H and $2K/n\mathrm{k}$ over the 10,000 trajectories at each time step during the evolution until the empirical reduced temperature $\theta^* := \mathrm{k}\theta/\epsilon$ reached the value 0.9. The results, plotted in Figure 1, show $\mathbf{E}^t\{H\}$ varying linearly with θ, Three points from the graph of U against T^*, obtained from simulations using the Gibbs equilibrium ensemble, are shown on the same diagram; the fact that they lie close to the graph of $\mathbf{E}^t\{H\}$ against θ provides evidence supporting Assumption 1.

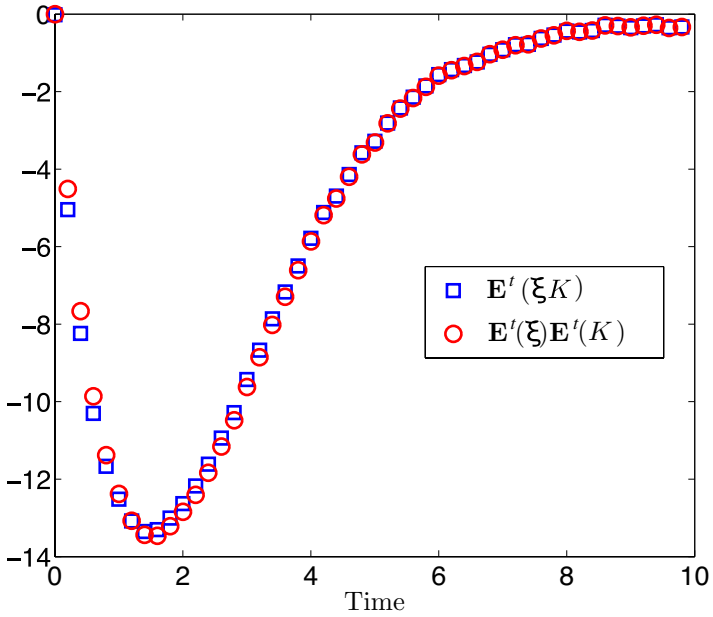

Fig. 2 Computed estimates of $\mathbf{E}^t\{\xi K\}$ and of $\mathbf{E}^t\{\xi\}\mathbf{E}^t\{K\}$ (in reduced units), plotted against time. As predicted by Assumption 2, the two curves are almost identical. (color online)

A separate calculation of the heat capacity, using the Gibbs fluctuation formula $C(T) = \mathrm{d}U(T)/\mathrm{d}T = (1/\mathrm{k}T^2)\mathbf{E}_{\mathrm{eq}}\left((H - \mathbf{E}_{\mathrm{eq}}(H))^2\right)$ gave the values $C(T) = 209.0, 208.4, 211.4$ at reduced temperatures $T^* = 0.6, 0.7, 0.8$. As predicted by Assumption 1, these numbers agree approximately with the slope of the graph, which is about 200.

Assumption 2 asserts that in the NHL thermostat the random variables ξ and K are uncorrelated (*i.e.* that $\mathbf{E}^t\{\xi K\} = \mathbf{E}^t\{\xi\}\mathbf{E}^t\{K\}$) for all t. Using the same sample of trajectories as in the previous test, we estimated the relevant expectations as the sample means of $\xi K, \xi$ and K at a succession of times t. The estimates of $\mathbf{E}^t\{\xi K\}$ and $\mathbf{E}^t\{K\}\mathbf{E}^t\{\xi\}$ are plotted against t in Figure 2. The close agreement of the two curves indicates that Assumption 2 is a good approximation.

6.2 Convergence to the thermostat temperature

To test the conclusions about convergence rates reached in Section 3, we used some new samples of 10,000 trajectories (still for the 108-particle Lennard-Jones system). For each sample the initial states were chosen at random using the Gibbs probability distribution (4) at (reduced) temperature $T_0^* = 1$

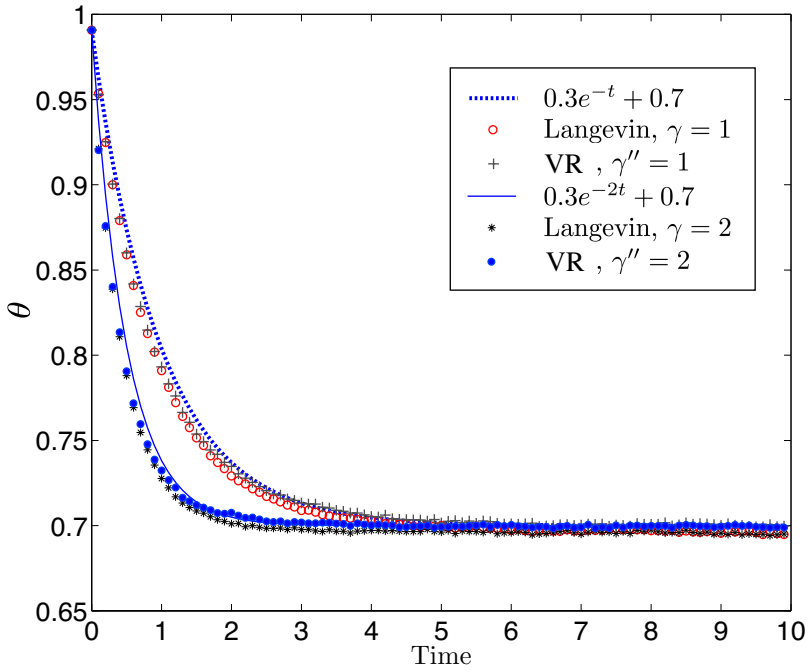

Fig. 3 Relaxation of the empirical temperature θ calculated as the mean of $2K/(nk)$ over a sample of 10,000 trajectories, both for the Langevin thermostat and the stochastic velocity rescaling thermostat (VR). The graph confirms that θ approaches the target temperature with exponential rate γ for the Langevin thermostat and γ'' for velocity rescaling. (color online)

and the temperature of the thermostat was $T^* = 0.7$, but different thermostats and different values for the parameters γ, etc. were used for the subsequent evolution.

6.2.1 Langevin and velocity rescaling thermostats

According to the theory in Section 3, both Langevin and the stochastic velocity rescaling method give exponential decay to equilibrium with the rates given in Table I. We tested this prediction for the Langevin thermostat with two samples of 10,000 trajectories using two different values of γ, computing the empirical temperature θ at various values of t as the sample mean of $2K/n$k and comparing with the theoretical prediction $\theta^*(t) = 0.7 + 0.3\mathrm{e}^{-\gamma t}$ given in equation (25). The results, shown in Figure 3, show that the calculated results are quite well represented by the theoretical curve. Figure 3 also shows the results of a similar test of the velocity rescaling thermostat, using the formula (11) for $\Phi(K)$ with two different values for γ''; again the theory agrees quite well with the simulation results.

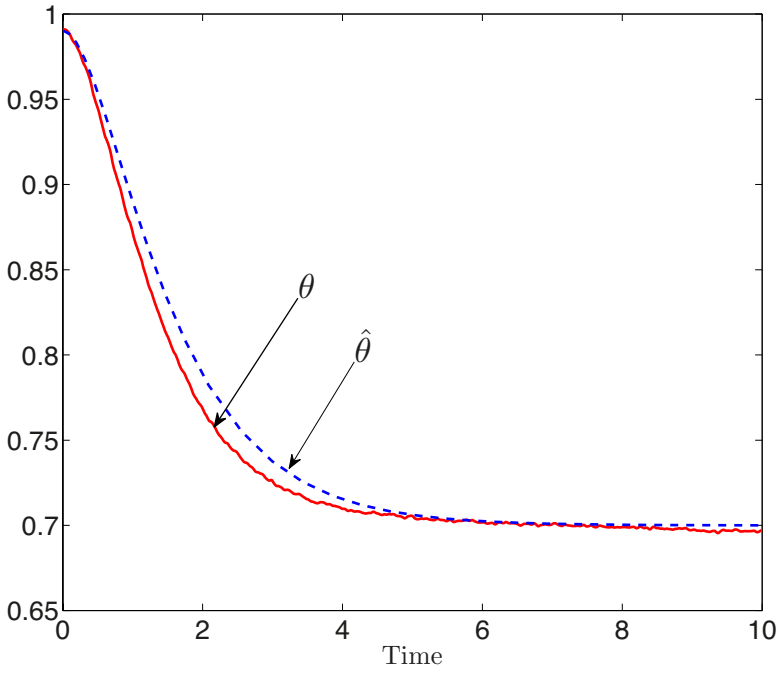

Fig. 4 Comparison of the relaxation of θ, obtained as the mean of $2K/nk$ over a sample of 10,000 NHL trajectories, with $\hat{\theta}$, the numerical solution of the nonlinear system (37), (38). (color online)

6.2.2 The NHL thermostat

In the case of the NHL thermostat, the approach to equilibrium is described in section 3 by the second-order system (37), (38) and therefore, according to the theory, neither of the variables $\theta, \mathbf{E}^t\{H\}$ decays as a simple exponential to its equilibrium value, even in the linear approximation. The best theoretical prediction about the decay of θ to its equilibrium value is obtained by solving the nonlinear equations (37), (38) numerically. Figure 4 shows such a prediction compared with values for θ obtained by averaging over 10,000 trajectories.

Since the system of differential equations (37), (38) describing the approach to equilibrium is now of second order, we also compared the phase portrait of that system — i.e. the trajectory of a point in the plane with coordinates $(\mathbf{E}^t\{\xi\}, \theta)$ computed from the second-order system— with the phase portrait of the sample means of ξ, θ at a succession of times t. The results are shown in Figure 5 and again the agreement is quite good.

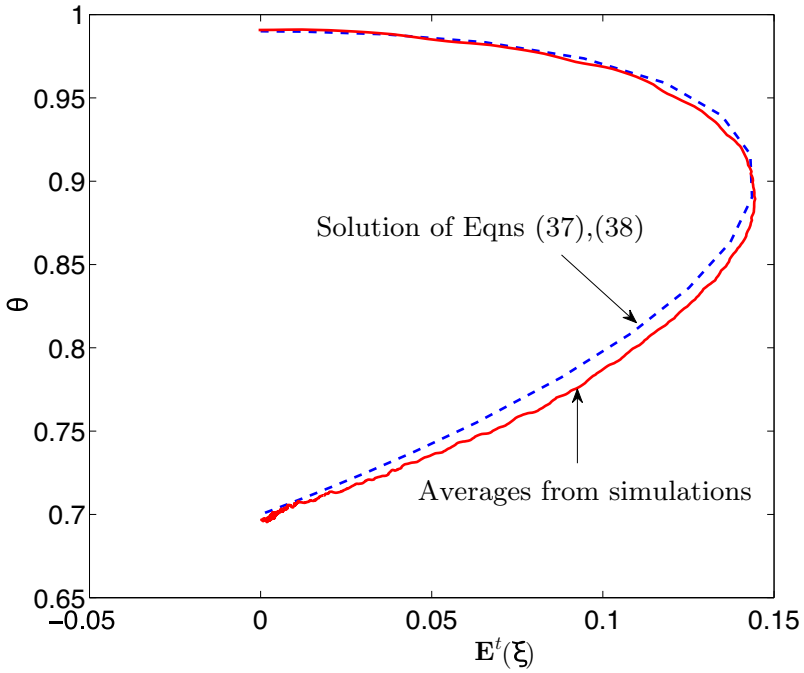

Fig. 5 Comparison of the curve $(\mathbf{E}^t\{\xi\}, \theta)$ obtained as the means of $2K/nk$ and ξ over a sample of 10,000 NHL trajectories, with the phase portrait of the nonlinear system (37), (38). (color online)

6.3 Growth of perturbations

We use the velocity autocorrelation function (VAF) to quantify the disturbance on the Hamiltonian dynamics. Figures 6, 7 show the errors $F_{LD} - F$ and $F_{NHL} - F$ against time. To calculate these errors, we first set up an ensemble of $N = 100,000$ equilibrated initial conditions $x^{(i)}(0), i = 1, \ldots, N$ at temperature $T^* = 0.7$. Next we computed $F(t)$, $F_{LD}(t), F_{NHL}(t)$ and $F_{VR}(t)$ by ensemble averaging; computing the evolution $x^{(i)}(t)$ using Hamiltonian dynamics, Langevin dynamics, NHL dynamics and stochastic velocity rescaling (VR) dynamics respectively and then approximating the VAFs by

$$F(t) \approx \frac{1}{NmnkT} \sum_{i=1}^{N} \sum_{j=1}^{n} p_j^{(i)}(0) p_j^{(i)}(t),$$

for a system of equal masses $m_j \equiv m$. When Langevin dynamics is used (Figure 6) we see that the error in the VAF grows linearly at first, the slope of the graph being the negative of the coupling coefficient γ which is also, in this case, the convergence rate. On the other hand for NHL (Figure 7), the error grows quadratically at first. Hence even for fast convergence rate (i.e. large γ') the VAF

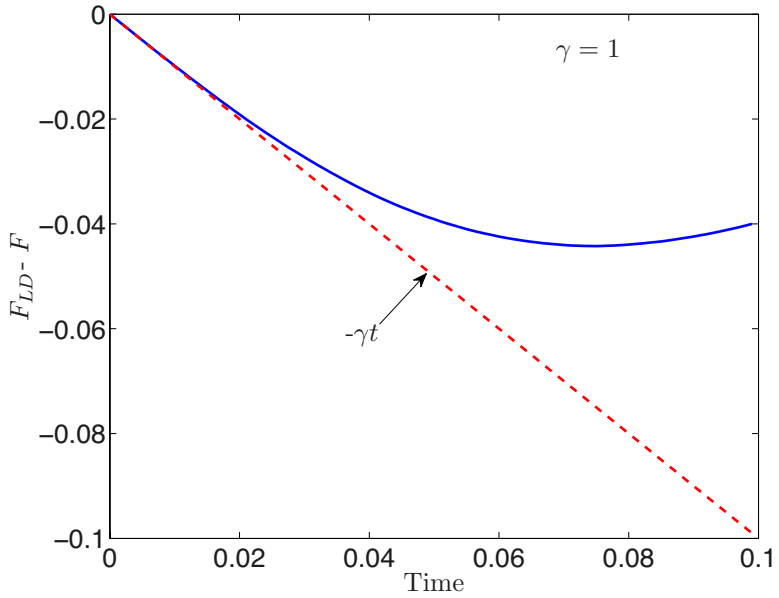

Fig. 6 The graph shows the computed perturbation in the VAF by Langevin dynamics (LD) for the 108 atom Lennard-Jones system, which clearly grows linearly (like $-\gamma t$), as anticipated (see Section 4.1). (color online)

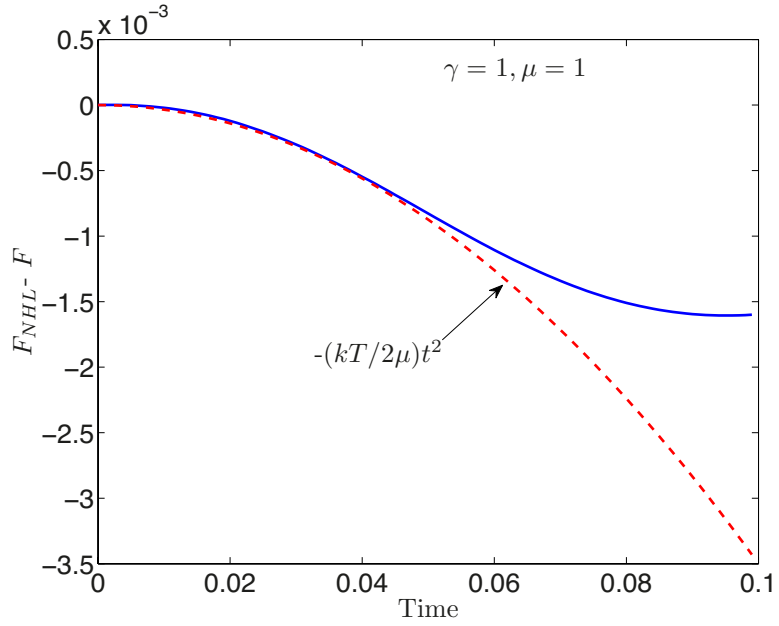

Fig. 7 For small times, the developing perturbation in the VAF by Nosé-Hoover-Langevin (applied to the 108 atom system) is well approximated by the quadratic derived in Section 4.3 (as $-\frac{kT}{2\mu}t^2$). (color online)

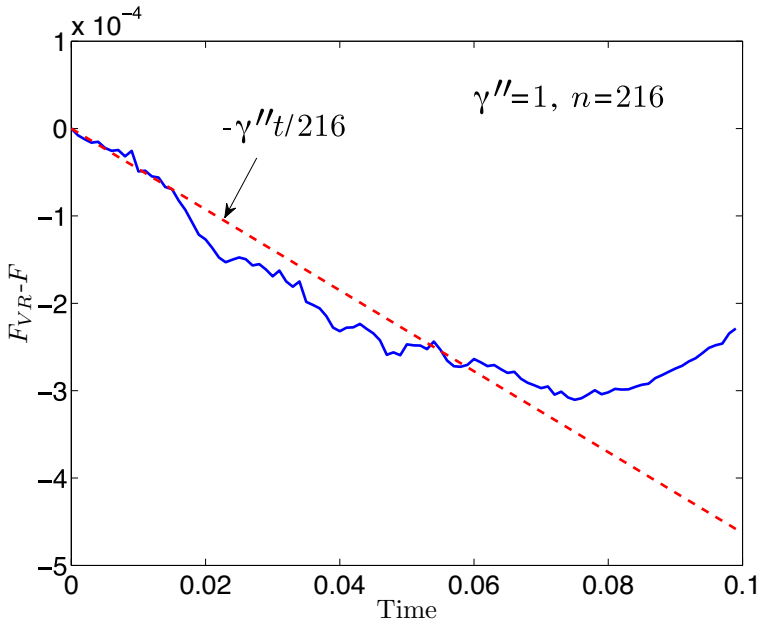

Fig. 8 The early evolution of the perturbation of the stochastic velocity rescaling thermostat for the 108 atom Lennard-Jones system . The calculation is impeded by the presence of multiplicative noise. (color online)

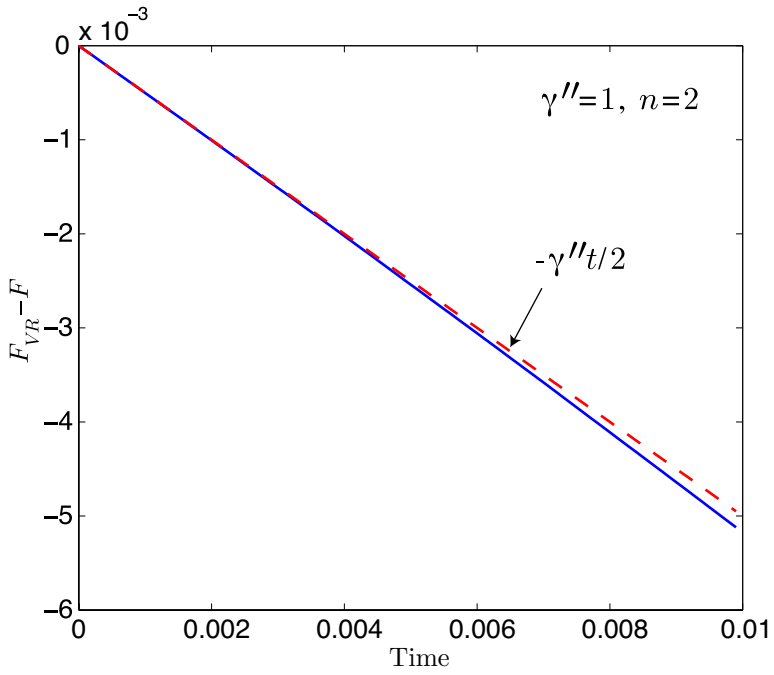

Fig. 9 The VAF perturbation for the stochastic velocity rescaling thermostat using a simple reduced 2-body Lennard-Jones model , on a short time interval. (color online)

calculated using the NHL method is much more accurate than that given by Langevin dynamics. Note the dramatic difference in scale between the perturbation observed here and that shown for Langevin dynamics in Figure 6.

We also examined the growth of the error in the velocity autocorrelation function for the stochastic velocity rescaling method. We found it more challenging to verify the linear growth estimated in Section 4.2 in a Lennard-Jones model, probably because the coefficient of linear growth is very small and the presence of multiplicative noise complicates the numerical calculation considerably (although Bussi et al [1,2] appear to have used an exact solution of the stochastic differential equations available for their specific choice of Φ). The result of our simulation is shown in Figure 8. When we re-ran the calculation using a simple two degree of freedom reduced Lennard Jones model with potential $U(q_1, q_2) = \varphi_{\mathrm{LJ}}(2q_1) + 2\varphi_{\mathrm{LJ}}(\sqrt{q_1^2 + q_2^2})$, a smaller stepsize, and computing more samples, the linear growth of VAF perturbation proved to be more easily verified, with the result shown in Figure 9.

6.4 Behavior of the Nosé-Hoover Langevin thermostat on longer time intervals

In our analysis we have relied on an expansion of the error in the velocity autocorrelation function which is only valid for small times; this is enough to distinguish the methods clearly in terms of their relative efficiencies, but it is interesting to ask if the same behavior carries over in practice to calculations of correlation functions on longer time intervals. To examine this question, we compared the errors in the velocity autocorrelation function computed using the Langevin and NHL methods. As determined analytically in Section 3 and numerically verified above, the rates of convergence of these methods are γ and $\gamma'/2$, respectively, where γ and γ' are the friction coefficients employed in the respective methods, as long as we remain below critical dampling. We solved the Lennard-Jones system of 108 atoms using the two methods, and using a coefficient for NHL which was twice that used for Langevin dynamics, in order to match the convergence rates; with the value $\mu = 1$ used here we were well below the critical damping threshold for NHL. Then we graphed the error in velocity autocorrelation functions for the different schemes. The results are show in Figure 10. This simulation demonstrates that the NHL method appears to give much smaller error in velocity autocorrelation function than does Langevin dynamics, when the collision parameters of NHL and Langevin dynamics are chosen

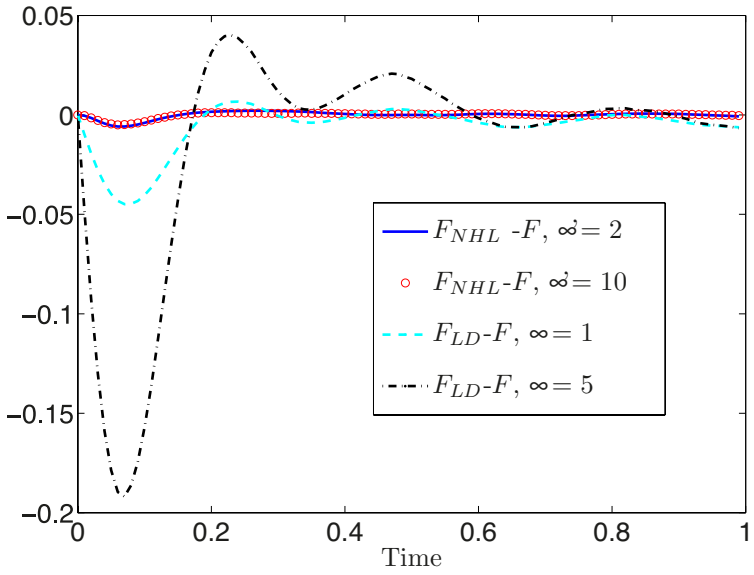

Fig. 10 Errors in velocity autocorrelation functions using Langevin and NHL dynamics, coefficients of each method chosen to give matching convergence rate of kinetic energy. Simulations for convergence rates of $r = 1$ and $r = 5$ are shown for each method. (color online)

in order to give matching kinetic convergence rates. Moreover, the graph in Figure 10 confirms the observation made in Section 4 that the perturbation of the velocity autocorrelation function in NHL is independent of γ' for short times (the $\gamma' = 2$ and $\gamma' = 10$ NHL curves appear to directly coincide in the early going).

7 Conclusions

In this article we have compared several methods for temperature control in molecular simulation in terms of their efficiencies. The efficiency is defined as the reciprocal of the magnitude of the perturbation of the velocity autocorrelation function incurred during the characteristic time it takes for the kinetic energy to converge to its equilibrium average. We obtain the convergence rate by an approximate treatment of the expectations of the stochastic differential equations describing the thermostat. An approximation to the growth rate of the perturbations is calculated by means of a Maclaurin expansion of the velocity autocorrelation function, derived from the stochastic differential equations. We find that for systems with n degrees of freedom, the Nosé-Hoover-Langevin thermostat and the stochastic velocity rescaling method both are order n times more efficient than Langevin dynamics in the sense

defined here. Where sampling trajectories are to be used for the computation of averaged dynamics (e.g. velocity autocorrelation functions), or consequent calculations such as diffusion rates or stress tensors, there would appear to be a clear advantage to using one of the "gentle" thermostats described here.

In comparing the Nosé-Hoover Langevin and Bussi-Parinello thermostats, we believe that there are certain advantages to using the former method. First, in the NHL method, the noise enters additively. It is well known that this simplifies both the analysis of the stochastic differential equations and the implementation of accurate numerical methods. Second, there are some rigorous (for special cases) analytical results concerning the ergodicity of the NHL method based on hypoellipticity [16], and similar results have not yet been established for stochastic velocity rescaling.

References

1. Bussi, G., Donadio, D., Parrinello, M.: Canonical sampling through velocity rescaling. J. Chem. Phys. **126**(1), 014,101 (2007)
2. Bussi, G., Parrinello, M.: Stochastic thermostats: comparison of local and global schemes. Computer Physics Communications **179**, 26–29 (2008)
3. Cancés, E., Legoll, F., Stoltz, G.: Theoretical and numerical comparison of some sampling methods for molecular dynamics. M2AN **41**(2), 351–389 (2007)
4. Chae, K., Elvati, P., Violi, A.: Effect of molecular configuration on binary diffusion coefficients of linear alkanes. J. Phys. Chem. B **115**(3), 500–506 (2011)
5. Cottrell, D., Tupper, P.: Energy drift in molecular dynamics simulations. BIT Numerical Mathematics **47**, 507–523 (2007)
6. Davidchack, R.L.: Discretization errors in molecular dynamics simulations with deterministic and stochastic thermostats. J. Comput. Phys. **229**(24), 9323 – 9346 (2010)
7. Ford, G.W., Kac, M.: On the quantum Langevin equation. J. Statist. Phys. **46**(5-6), 803–810 (1987)
8. Garrido, P., Gallavotti, G.: Boundary dissipation in a driven hard disk system. J. Stat. Phys. **126**, 1201–1207 (2007)
9. Hoover, W.: Canonical dynamics: equilibrium phase space distributions. Phys. Rev. A. **31**, 1695–1697 (1985)
10. Hörmander, L.: Hypoelliptic second order differential equations. Acta Mathematica **119**(1), 147–171 (1967)
11. Hörmander, L.: The analysis of linear partial differential operators. Springer Verlag (1985)

12. Hünenberger, P.: Thermostat algorithms for molecular dynamics simulations. Adv. in Poly. Sci. **173**, 105–149 (2005)
13. Langevin, P.: On the Theory of Brownian Motion (Sur la thórie du mouvement Brownien). C. R. Acad. Sci. (Paris) **146**, 530–533 (1908)
14. Legoll, F., Luskin, M., Moeckel, R.: Non-ergodicity of the Nose-Hoover thermostatted harmonic oscillator. Arch. Rational Mech. Anal. **184**, 449–463 (2007)
15. Legoll, F., Luskin, M., Moeckel, R.: Non-ergodicity of Nosé-Hoover dynamics. Nonlinearity **22**, 1673–1694 (2009)
16. Leimkuhler, B., Noorizadeh, N., Theil, F.: A gentle stochastic thermostat for molecular dynamics. J. Stat. Phys. **135**(2), 261–277 (2009)
17. Mattingly, J.C., Stuart, A.M., Higham, D.J.: Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. Stochastic Process. Appl. **101**(2), 185–232 (2002)
18. Nosé, S.: A unified formulation of the constant temperature molecular dynamics method. J. Chem. Phys. **81**, 511–519 (1984)
19. Oksendal, B.K.: Stochastic Differential Equations: An Introduction with Applications, 4th edn. Springer (1995)
20. Samoletov, A., Chaplain, M.A.J., Dettmann, C.P.: Thermostats for "slow" configurational modes. J. Stat. Phys. **128**, 1321–1336 (2007)
21. Sun, J., Zhanga, L.: Temperature control algorithms in dual control volume grand canonical molecular dynamics simulations of hydrogen diffusion in palladium. J. Chem. Phys. **127**, 164,721 (2007)
22. Williams, G.: Time-correlation functions and molecular motion. Chem. Soc. Rev. **7**(89), 89–131 (1978)
23. Yamashita, H., Endo, S., Wako, H., Kidera, A.: Sampling efficiency of molecular dynamics and monte carlo method in protein simulation. Chemical Physics Letters **342**(3-4), 382 – 386 (2001)